

POWERED BY **Dialog**

Number of Patents: 8 Number of Countries: 4 Number of Legal Status Entries: 15
 Patent Basic (No,Kind,Date): EP 1063831 A2 20001227

Network status server, information distribution system, control method, and storage medium for storing control program (English)

Netzwerkzustandsserver, Informationsverteilungssystem, Steuerverfahren, und Speichermedium zur Speicherung eines Steuerungsprogramms (German)

Serveur d'etat de reseau, systeme de distribution d'information, methode de controle, et support d'enregistrement pour enregistrer un programme de controle (French)

Patent Assignee: CANON KK (JP)

Author (Inventor): OGINO TSUKASA (JP)

Record Type: Legal Status; Abstract; Cited Refs

Patent Family

Patent Number	Kind	Date	Application Number	Kind	Date	Update
CN 1280342	A	20010117	CN 2000124151	A	20000623	200104
CN 1153430	C	20040609	CN 2000124151	A	20000623	200513
DE 60030508	D1	20061019	DE 60030508	A	20000623	200641
DE 60030508	T2	20070516	DE 60030508	T	20000623	200720
EP 1063831	A2	20001227	EP 2000305308	A	20000623	200051 (B)
EP 1063831	A3	20040211	EP 2000305308	A	20000623	200406
EP 1063831	B1	20060906	EP 2000305308	A	20000623	200633
JP 2001007844	A	20010112	JP 1999178474	A	19990624	200108

Priority Data

Application Number	Kind	Date
JP1999178474	A	19990624

CHINA (CN)

China (CN) Patent(s):

Patent (Number,Kind,Date): CN 1280342 A 20010117

Network state service device, information distribution system, control method and storage medium (English)

Patent Assignee: CANON KK (JP)

Author (Inventor): TSUKASA OGINO (JP)

Priority (Number,Kind,Date): JP 1999178474 A 19990624

Applic (Number,Kind,Date): CN 2000124151 A 20000623

ECLA: H04L-029/08N9A

IPC + Level Value Position Status Version Action Source Office

v. 7 main: G06F-015/16

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 1 063 831 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
27.12.2000 Bulletin 2000/52

(51) Int Cl.7: H04L 29/06

(21) Application number: 00305308.9

(22) Date of filing: 23.06.2000

(84) Designated Contracting States:
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: 24.06.1999 JP 17847499

(71) Applicant: CANON KABUSHIKI KAISHA
Tokyo (JP)

(72) Inventor: Ogino, Tsukasa,
Canon Kabushiki Kaisha
Tokyo (JP)

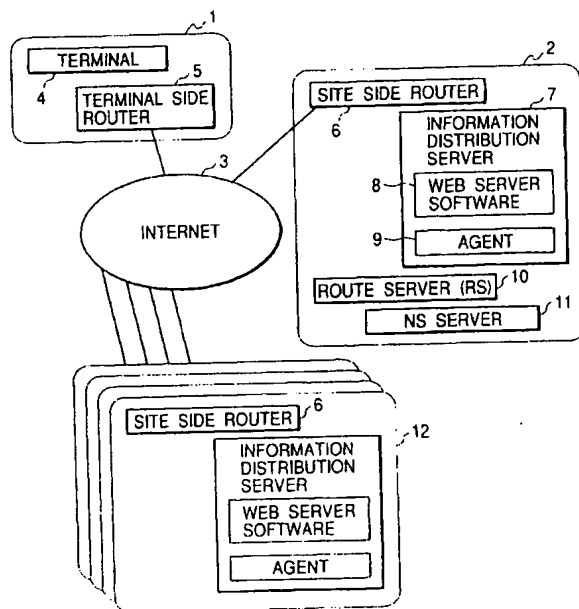
(74) Representative:
Beresford, Keith Denis Lewis et al
BERESFORD & Co.
High Holborn
2-5 Warwick Court
London WC1R 5DJ (GB)

(54) **Network status server, information distribution system, control method, and storage medium for storing control program**

(57) There is disclosed a network status server for receiving an inquiry from a web server accessed by a client, and using path information and determining the optimum web server to return a response to the accessed web server. The web server receiving the re-

sponse instructs the client to access the optimum web server from the next time. In this constitution, accesses among the dispersed/arranged web server sites are leveled, and access loads can be leveled even among the web servers in the respective web server sites. Moreover, the client can access the optimum web server.

FIG. 1



EP 1 063 831 A2

Description

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] The present invention relates to a network status server for distributing large-capacity data on a network, an information distribution system, a control method, and a storage medium for storing a control program.

[0002] Particularly, the present invention relates to a network status server, an information distribution system, a control method and a storage medium for storing a control program when a plurality of dispersed/arranged information distribution servers are controlled.

Related Background Art

[0003] As one of information distribution techniques using Internet, World Wide Web (hereinafter referred to as Web) has generally been used. This technique comprises: operating a web server software on a general computer (information processor); and distributing information such as contents (characters, charts, tables, and the like) described in an HTML language. The computer as a client which accesses a server (hereinafter referred to as the access client) may be connected to the server via a network, and free perusal is possible with a general-purpose perusal software (browser). In recent years, a web server site accessible to the entire world has appeared and the importance has further increased. On the other hand, when the web server has been accessed from the entire world in a concentrated manner, a large load is applied to the connection network, and the web server. To solve the problem, several techniques are currently proposed as countermeasures. The techniques will briefly be described with reference to Figs. 6 and 7.

[0004] First, a system in which a virtual host computer is utilized as shown in Fig. 6 is proposed as a technique of dispersing accesses concentrated to the web server. This aims at dispersing the above-described load to a plurality of servers. In this example, it is supposed that the access client accesses URL (access address) of www.abcd.jp. In the access to www.abcd.jp, first the virtual host computer is accessed. The virtual host computer intends to appropriately disperse the accesses from the load situation of the respective subordinate web servers. Actually, any one of www1, www2, www3 distributes predetermined information in response to the access from the client, but the virtual host computer once accepts the access to appropriately disperse the access. Moreover, a transfer system by TCP connection hop is proposed. The system will be described with reference to Fig. 7. When there is an access from the client, first www1 is accessed. In the www1, a scheduler operates, and selects the optimum web server www3, and responds to the access from the client. Specifically, the www1 constantly responds to an access request, and the optimum web server (www1, www2, www3) responds. In this system, the access dispersion is intended in the schedule function.

[0005] However, the above-described conventional access dispersion system has the following problems.

1) When the virtual host computer shown in Fig. 6 is utilized, a problem occurs that a device serving as an exclusive virtual host computer is separately necessary.

2) Moreover, when the exclusive virtual host computer temporarily accepts all the accesses, and when there are many accesses, the processing ability of the virtual host computer raises a problem. Specifically, the processing ability of the virtual host computer substantially indicates the processing ability of the web server site (the site in which the web server operates).

3) Moreover, in the virtual host computer system, the access is accepted in the server site in which the virtual host computer is present, the access of the web server in its own server site is leveled, and the respective virtual host computers are necessary for the respective web server sites provided with the dispersed/arranged web servers. Moreover, predetermined different URLs need to be set in the respective virtual host computers. Therefore, a problem occurs that the accessing client has to select the dispersed/arranged web server site by itself.

4) Furthermore, since the accessing client selects the web server site, a problem occurs that the accesses to the respective web server sites cannot be leveled.

However, the transfer system by the TCP connection hop as shown in Fig. 7 has an advantage that, different from the virtual host computer system, a mechanism for temporarily accepting the exclusive access is not separately necessary.

5) However, the accesses are concentrated on the web server in which the scheduler is permanently stationed. Therefore, if a trouble is generated in the web server in which the scheduler is permanently stationed, a problem occurs that the entire server site cannot function.

6) Moreover, in this system, the access to the dispersed/arranged web server site cannot be controlled to the optimum. Since the system is constituted for a purpose of leveling the access in the web server site (in its own

network) managed by the scheduler, predetermined different URLs need to be set in the respective web server sites similarly as the above-described virtual site system. Therefore, there is a problem that the accessing client selects the dispersed/arranged web server site by itself.

7) Furthermore, similarly as the above-described virtual site system, since the accessing client selects the web server site, the accesses to the respective web server sites cannot be leveled.

SUMMARY OF THE INVENTION

[0006] The present invention is constituted by a network status server comprising: logical distance obtaining means for obtaining respective logical distances between respective sites provided with dispersed/arranged information distribution servers and an accessed client; and site determination means for determining an optimum site from the respective sites provided with the dispersed/arranged information distribution servers based on the obtained logical distances.

[0007] Moreover, the present invention is constituted by a network status server comprising: logical distance obtaining means for obtaining respective logical distances between respective sites provided with dispersed/arranged information distribution servers and an accessed client; collection means for collecting network state information between the respective sites provided with the dispersed/arranged information distribution servers and the accessed client, and state information in the site provided with the dispersed/arranged information distribution servers; and server determination means for determining the optimum information distribution server from the dispersed/arranged information distribution servers based on the logical distance obtained by the logical distance obtaining means, the network state information between the respective site provided with the dispersed/arranged information distribution servers and the accessed client, and the state information in the sites provided with the dispersed/arranged information distribution servers collected by the collection means.

[0008] Furthermore, the present invention is constituted by a network status server comprising: logical distance obtaining means for obtaining respective logical distances between respective sites provided with dispersed/arranged information distribution servers and an accessed client; collection means for collecting network state information between the respective sites provided with the dispersed/arranged information distribution servers and the accessed client, and state information in the site provided with the dispersed/arranged information distribution servers; site determination means for, when it is judged that the access from the client is a first access, determining an optimum site from the respective sites provided with the dispersed/arranged information distribution servers based on the logical distance obtained by the logical distance obtaining means; first server determination means for determining an optimum information distribution server from the information distribution servers in the optimum site determined by the site determination means based on a predetermined condition; and second server determination means for, when it is judged that the access from the client is not the first access, determining the optimum information distribution server from the dispersed/arranged information distribution servers based on the logical distance obtained by the logical distance obtaining means, the network state information between the respective sites provided with the dispersed/arranged information distribution servers and the accessed client, and the state information in the site provided with the dispersed/arranged information distribution servers collected by the collection means.

[0009] Moreover, according to the present invention, there is provided an information distribution system comprising: dispersed/arranged information distribution servers; and a network status server.

[0010] The information distribution server comprises: inquiry means for inquiring of the network status server about the optimum information distribution server among the dispersed/arranged information distribution servers in response to an accessing client; response reception means for receiving a response from the network status server; and indication means for indicating the optimum information distribution server to the accessing client. The network status server comprises: inquiry reception means for receiving the inquiry from the inquiry means; logical distance obtaining means for obtaining respective logical distances between respective sites provided with the dispersed/arranged information distribution servers and the accessed client; site determination means for determining the optimum site from the respective sites provided with the dispersed/arranged information distribution servers based on the obtained logical distance; server determination means for determining the optimum information distribution server from the information distribution servers in the optimum site determined by the site determination means based on a predetermined condition; and response transmission means for returning the server determined by the server determination means as the optimum server to the information distribution server having transmitted the inquiry.

[0011] Furthermore, according to the present invention, there is provided an information distribution system comprising: dispersed/arranged information distribution servers; and a network status server.

[0012] The information distribution server comprises: inquiry means for inquiring of the network status server about the optimum information distribution server among the dispersed/arranged information distribution servers in response to an accessing client; response reception means for receiving a response from the network status server; and indication means for indicating the optimum information distribution server to the accessing client. The network status

server comprises: inquiry reception means for receiving the inquiry from the inquiry means; logical distance obtaining means for obtaining respective logical distances between respective sites provided with the dispersed/arranged information distribution servers and the accessed client; collection means for collecting network state information between the respective sites provided with the dispersed/arranged information distribution servers and the accessed client, and state information in the site provided with the dispersed/arranged information distribution servers; server determination means for determining the optimum information distribution server from the dispersed/arranged information distribution servers based on the logical distance obtained by the logical distance obtaining means, the network state information between the respective sites provided with the dispersed/arranged information distribution servers and the accessed client, and the state information in the site provided with the dispersed/arranged information distribution servers collected by the collection means; and response transmission means for returning the server determined by the server determination means as the optimum server to the information distribution server having transmitted the inquiry.

[0013] Additionally, according to the present invention, there is provided an information distribution system comprising: dispersed/arranged information distribution servers; and a network status server.

[0014] The information distribution server comprises: inquiry means for inquiring of the network status server about the optimum information distribution server among the dispersed/arranged information distribution servers in response to an accessing client; response reception means for receiving a response from the network status server; and indication means for indicating the optimum information distribution server to the accessing client. The network status server comprises: inquiry reception means for receiving the inquiry from the inquiry means; logical distance obtaining means for obtaining respective logical distances between respective sites provided with the dispersed/arranged information distribution servers and the client accessing the information distribution server; collection means for collecting network state information between the respective sites provided with the dispersed/arranged information distribution servers and the accessed client, and state information in the site provided with the dispersed/arranged information distribution servers; site determination means for, when it is judged that the access from the client is a first access, determining the optimum site from the respective sites provided with the dispersed/arranged information distribution servers based on the logical distance obtained by the logical distance obtaining means; first server determination means for determining the optimum information distribution server from the information distribution servers in the optimum site determined by the site determination means based on a predetermined condition; second server determination means for, when it is judged that the access from the client is not the first access, determining the optimum information distribution server from the dispersed/arranged information distribution servers based on the logical distance obtained by the logical distance obtaining means, the information of the network state between the respective sites provided with the dispersed/arranged information distribution servers and the accessed client collected by the collection means, and the state information in the site provided with the dispersed/arranged information distribution servers; and response transmission means for returning the server determined by the first server determination means or the second server determination means as the optimum server to the information distribution server having transmitted the inquiry.

[0015] The present invention relates to the server, the system, a control method of controlling the server and system, or a storage medium for storing a control program for controlling the server and system.

[0016] Since the present invention is constituted as described above, the access load between the dispersed/arranged web server sites can be leveled.

[0017] Moreover, the access between the dispersed/arranged web server sites is leveled, while the access load between the web servers in each web server site can be leveled.

[0018] Furthermore, in order to determine the optimum web server at a high speed, the optimum web server site is first determined at a high speed by a first determination method, and the true optimum web server is then determined by a second determination method, so that the optimum web server can be determined at the speed as high as possible.

[0019] Additionally, in order to determine the optimum web server at a high speed, the accessed web server can detect the network state and server state at a high speed.

[0020] Moreover, the access client can be switched to the optimum web server without requiring any active operation.

[0021] Furthermore, when all the dispersed/arranged web servers accept the first access, the concentration of the accesses can be eliminated.

[0022] Other features and advantages of the patent invention will be apparent from the following description in conjunction with the accompanying drawings, in which like reference characters designate the same or similar parts throughout the figures thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

[0023] The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate embodiments of the invention and together with the description, serve to explain the principles of the invention.

[0024] Fig. 1 is an explanatory view of a constitution according to an embodiment of the present invention.

[0025] Fig. 2 is an explanatory view of a function and operation according to the embodiment of the present invention.

[0026] Figs. 3A, 3B, 3C, 3D and 3E are flowcharts showing the function and operation according to the embodiment of the present invention.

[0027] Fig. 4 is an explanatory view showing the function and operation of an NS server according to the embodiment of the present invention.

[0028] Fig. 5 is an explanatory view showing the function and role according to the embodiment of the present invention.

[0029] Fig. 6 is an explanatory view of a conventional virtual host computer system.

[0030] Fig. 7 is an explanatory view of a conventional TCP connection hop system.

[0031] Fig. 8 is a diagram showing an approximation curve during approximation of collected information with a predetermined function.

[0032] Fig. 9 is a block diagram of a computer which can be realized by an NS server, information distribution server and route server of the present invention.

[0033] Fig. 10 is a schematic view showing a storage medium for storing a program code and a transmission device for transmitting the program code.

[0034] Fig. 11 is an image diagram showing the memory map of the storage medium for storing a network status server control program or an information distribution system control program in the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0035] An embodiment of the present invention will be described in detail hereinafter with reference to the drawings. Fig. 1 is a constitution system diagram of an optimum server selection system of the present invention. Shown is an example in which Internet is used as communication means between respective sites. In Fig. 1, a terminal side site 1 accesses an information distributing server, and a server site 2 is provided with an accessed information distributing server (hereinafter referred to as the information distribution server or the web server). Here, the site means the same logical network block, and indicates, for example, the address block of class B or C as a network address class. Of course, even when the block is physically in the same place, it can be regarded as the site. Moreover, in the present embodiment, the server site 2 is dispersed and disposed to perform high-speed large-capacity information distribution. Fig. 1 shows dispersed/arranged server sites 12. Numeral 3 denotes Internet as communication means among the respective sites. In the present embodiment, the communication means among the sites is Internet, but is not limited to this, and for example, a packet communication network, a public network (telephone network), and the like may be used.

[0036] A terminal 4 accesses the information distribution server, and is an information processor in which a standard browser accessible to the information distribution server (web server) is installed in the present embodiment. A terminal side router 5 is a router for performing the path control of the inside and outside of an access side site. A site side router 6 is a router for performing the path control of the inside and outside of the site provided with the information distribution server.

[0037] In an information distribution server 7, a web server software 8 as a standard information distributing software and an agent software 9 newly prepared in the present invention operate. The web server software 8 is a server software provided with a function of transferring information written in html language in response to a request from the outside. Moreover, the agent software 9 is a software provided with various server operation situations and indications for the web server software 8. Additionally, needless to say, the information distribution server may operate with the software integrally formed by the web server software 8 and agent software 9.

[0038] The respective site side routers 6 are provided with the path information obtained when the respective sites view the other sites through Internet (network). A route server 10 is provided with a function of obtaining path information (e.g., BGP information, RIP information, OSPF information, and the like) from the respective site side routers 6. Moreover, the route server 10 has a function of transmitting the path information to a network status server (hereinafter referred to as the NS server) 11.

[0039] The NS server 11 instructs the agent 9 to measure various operation states of the site 2 provided with the dispersed/arranged information distribution servers and the information distribution servers in the site 12, and obtains and stores measurement results. Moreover, the agent 9 measures the network states (e.g., congestion degree, router step number, packet loss ratio, response time, and the like) between terminal side site 1 having requested for the information and the respective information distribution servers 2 and 12, and the NS server 11 obtains and stores the network states measured by the agent 9. Furthermore, the NS server 11 obtains the information from the route server 10 to obtain the network distance between the terminal side site 1 having requested for the information and the sites 2 and 12 provided with the respective information distribution servers.

[0040] Additionally, in Fig. 1, the route server 10 and NS server 11 are in the site 2, but are not limited in the site 2, and may be in any other site.

[0041] An operation for controlling the access to the optimum web server in the present embodiment will next be described. First, Fig. 2 is a schematic view of the web server access. Here, an operation performed when the terminal 4 of the site 1 accesses the web server is concretely shown. Shown is a case in which the terminal 4 in the autonomous (AS) system 1 accesses the web server provided with the access address (hereinafter referred to as URL) of the web site of www.abcd.e.jp. First, in order to obtain the IP address of the target web server, a nearest domain name server (DNS) in its network (site 1) is inquired from the URL, and the corresponding IP address (addr1) is received (201). Subsequently, the desired web site (the web site having the URL of www.abcd.e.jp) is accessed (= requested) by the IP address (addr1) (202). When the web server A having the IP address (addr1) accepts the access, the agent 9 inquires of the network status server (NS server) whether or not the web server is an optimum server for the accessed client (203). The NS server inquires of the route server 10 about the network distance between the respective sites 2, 3, 4 provided with the dispersed/arranged web servers and the accessed client site 1 (204). Since the route server 10 exchanges the path information with the border gateway router 6 in each site, the distance between each site and the client site 1 can quickly be returned (205). The NS server 11 receives the response from the route server 10, determines the site 3 as the closest site in the network, and returns the URL of the optimum web server B to the agent 9 of the site 4 (206). Alternatively, the IP address (addr2) of the optimum web server may be returned. In the present embodiment, the AS path information of the border gateway protocol (BGP) is used as a site selection method at this time (when the client first accesses the site) so that the information can quickly be obtained. For this the site 3 is selected as the site with a short AS path from each web server to the client site, and a predetermined method is used to select the web server B having addr2 in the site 3 as the optimum web server.

[0042] Additionally, in addition to the path information, the NS server operates to obtain number of routers between each site and the client (router hop number), response time, network information such as a packet loss ratio, network information in each site (congestion degree, number of packets, number of packet errors, and the like), server load situation (CPU load ratio, CPU idle ratio, number of connection links, disk load ratio, and the like), and other information. The NS server prepares for the next inquiry and stores the information (including temporary storage) (207).

[0043] The agent 9 in the web server A instructs the client 4 to access the web server B during the next access (208). For example, the web server to be connected to the client can automatically be changed using the location command of the HTML language. As a result, the client 4 will automatically access the optimum web server B from the next access.

[0044] An operation performed when the client 4 continuously accesses the web server B will next be described similarly with reference to Fig. 2. The web server B is accessed through the above-described process (209). When the web server B accepts the access, the agent 9 inquires of the NS server 11 whether or not the web server itself is an optimum server for the accessed client (210). The NS server 11 collects the path information and integrates various state information similarly as the operation during the above-described first access and determines the optimum server site and optimum web server. In addition to the path information obtained and stored by the previous access, the NS server 11 obtains number of routers between each site and the client (number of router hops), response time, network information such as the packet loss ratio, network information in each site (congestion degree, number of packets, number of packet errors, and the like), and server state information (CPU load ratio, CPU idle ratio, number of connection links, disk load ratio, and the like) which could not be inspected during the previous access, and the optimum web server to be accessed by the client is calculated and determined from the obtained various information. At this time the selection method comprises applying a numeric weight to each information in accordance with the content, and comparing the total weight for each site or each server to determine the optimum web server C (211). As a result, the client 4 will automatically access the optimum web server C from the next access (214).

[0045] Furthermore, when the client 4 successively accesses the web server C, the operation similar to the above-described operation of accessing the web server B is performed (215, 216), but the operation of accessing the NS server is performed after a predetermined time. Subsequently, when a new optimum web server is determined by the obtained and stored various network state information and server state information in the predetermined time, the client 4 changes the access to the new web server.

[0046] An operation of controlling the access to the optimum web server of the present embodiment will next be described with reference to Figs. 3A to 3E. Fig. 3A is a flowchart showing an operation performed from when the access client first accesses the web server A until the access is changed to the first optimum web server B. Fig. 3B is a flowchart showing the operation of the NS server. Fig. 3C is a flowchart showing an operation performed from when the access client next accesses the web server B until the access is changed to the optimum access web server C. Fig. 3D is a flowchart showing the operation of the NS server. Fig. 3E is a flowchart showing an operation performed from when the access client accesses the web server C and further inspects the optimum access web server until the web server C is again accessed.

[0047] The operation will first be described with reference to Figs. 3A, 3B. In order to obtain the IP address of the target web server from the URL, the closest domain name server (hereinafter abbreviated as DNS) in its network (site 1) is inquired, and the corresponding IP address (addr1) is received (S1).

[0048] Subsequently, the desired web site A is accessed by the IP address (addr1) (S2).

[0049] When the web server A accepts the access, the agent 9 inquires of the NS server whether or not the web server is an optimum server for the accessed client (S3).

[0050] In the NS server, the network distance from the dispersed/arranged respective web server sites to the accessed client site 1 is inquired of the route server 10 (S4).

[0051] Since the route server 10 exchanges the path information with the site side router 6 in each site (hereinafter referred to as the border gateway router), the distance from each site to the client site 1 can quickly be returned, and the NS server obtains the path information and distance (S5).

[0052] The NS server 11 receives the response from the route server 10, fixes the site 3 as the close site in the network, determines the optimum web server B in the site 3 in a predetermined method, and answers the agent 9 of the site 4 that the optimum web server is the web server B (S6). Additionally, the predetermined method of fixing the optimum web server in S6 may comprise predetermining the web server for each site, determining the web server in a predetermined order for each site, or arbitrarily determining the web server in the site. Moreover, since the NS server obtains the state information in the site provided with the web server (e.g., the information of the network state in the site such as the congestion degree, number of packets and packet errors, and the server state information such as CPU load ratio, CPU idle ratio, number of connection links and disk load ratios), the optimum web server may be determined utilizing the state information in the site.

[0053] The NS server stores the obtained path information (including the temporary storage) to prepare for the next inquiry (S7).

[0054] On the other hand, the agent 9 in the web server A answers the client 4 that the optimum web server B is accessed from the next access (S8).

[0055] As a result, the client 4 will automatically access the optimum web server B from the next access.

[0056] The operation will next be described with reference to Figs. 3C, 3D.

[0057] The web server B is accessed through the above-described process (S9).

[0058] When the web server B accepts the access, the agent 9 inquires of the NS server 11 whether or not the web server is an optimum server for the accessed client (S10).

[0059] The NS server 11 collects the path information and integrates various state information similarly as the operation during the above-described first access and determines the optimum server site and optimum web server. In addition to the path information obtained and stored by the previous access, the NS server 11 obtains number of routers between each site and the client (router hop number), response time, network information such as the packet loss ratio, network information in each site (congestion degree, number of packets, packet error, and the like), and web server state information (CPU load ratio, CPU idle ratio, number of connection links, disk load ratio, and the like) which could not be inspected during the previous access, and the optimum web server to be accessed by the client is calculated and determined from the obtained various information. At this time the selection method comprises applying the numeric weight to each information in accordance with the content, and comparing the total weight for each site or each server to determine the optimum web server C (S11).

[0060] The NS server 11 notifies the agent 9 in the web server B of the optimum web server C (S12), and further answers the client 4 that the optimum web server C is accessed from the next access (S13).

[0061] As a result, the client 4 will automatically access the optimum web server C from the next access (S14).

[0062] Finally, the operation will be described with reference to Fig. 3E. When the client 4 successively accesses the web server C, the optimum web server is inquired of the NS server similarly as the above-described access to the web server B (S15), but here the web server C is judged to be the optimum web server similarly as the previous access, and the processing for changing the web server to be accessed (S16) is not performed. However, when it is answered in S15 that the web server different from the web server C is the optimum web server, needless to say the processing for changing the web server to be accessed is performed.

[0063] Additionally, the operation of accessing the NS server is intermittently performed after a predetermined time. Subsequently, when a new optimum web server is determined by the obtained and stored various network state information and server state information in the predetermined time, the client 4 changes the access to the new web server.

[0064] An operation will next be described with reference to Fig. 4, in which the NS server 11 obtains and stores the state of the network from the web server to the client, the state of the network in the web server site, and information on the web server state.

[0065] In the present system, the following two categories of information are periodically (intermittently) collected:

- (1) Network state from each server site to the client;
- (2) State in each server site (network state, server state).

[0066] Fig. 4 shows a relation. The NS server 11 periodically measures the network state from each server site to the accessed client. Items to be measured are described later.

[0067] Each server site router 6 holds the network path information seen from the server site, and the route server

10 collects the path information owned by the router from each server site router 6. The route server 10 has a logical connection as being peer-to-peer with each server site router 6, that is, the border gateway router (BGR) or an external network border router (401). Subsequently, the analysis result of the predetermined path information is obtained in response to the request from the NS server 11, and returned to the NS server 11 (402). For other inspection items, the NS server autonomously performs the inspection (403), but in some cases the NS server 11 instructs the agent software 9 incorporated in each web server software to perform inspection in accordance with the inspection content.

[0068] For the network and server states in each server site, the agent software 9 incorporated in the web server software of each server site periodically performs the inspection in accordance with the instruction from the NS server 11. The collected various information are stored in the NS server in time series. In this data storage method, time is intermittently stored as information attached to the various information. In the present embodiment, various information are inspected, obtained, and stored every 30 minutes. Moreover, it is empirically known that the stored information periodically changes in predetermined periods such as a day unit, week unit, month unit, and year unit. Therefore, in the present system, the stored information is approximated with a predetermined function.

[0069] The time unit for the approximation may be any one of the day unit, week unit, month unit, and year unit, but in the present embodiment, various information are approximated by the day unit by a three-dimensional function and a coefficient is saved. For example, when the congestion degree (CSn) of the network in the server site is approximated with the three-dimensional function, a time axis is set to t, and the approximation is performed with an approximation curve shown in Fig. 8. When the approximation equation is represented as $CSn = a \cdot t^3 + b \cdot t^2 + c \cdot t + d$ (a, b, c, d being coefficients), the NS server saves this coefficient (the three-dimensional function itself may be saved).

[0070] As a result, when the request for selecting the optimum server is received from each server, by utilizing the three-dimensional function, various states are momentarily estimated, and the optimum server is selected based on the estimated value.

[0071] The collection of the network state information between each server site and the access client is not infinitely continued, and is continued while the access to the dispersed/arranged servers is continued and for a predetermined period after that. In the present embodiment, the period is set to one week. When there is no access from the access client 4 for one week, the information collecting operation to the access client 4 from each site is not performed.

[0072] Moreover, when there is no access from the access client 4 for the predetermined period (e.g., three weeks), the network state information (the approximation equation, and the like) between the access client 4 and each server site stored in the NS server stored in the NS server may be discarded. In this case, no wasteful information is stored, and the storage capacity necessary for the NS server can advantageously be saved.

[0073] Furthermore, when there is no access from the access client 4 for the predetermined period, the path information between the access client 4 and each server site stored in the NS server may also be discarded.

[0074] Additionally, the stored information can also be treated as the result representative of a specific area. Specifically, the NS server obtains and stores the path information (logical distance) between each server site and the access client or the network state information between each server site and the access client. For the path information (logical distance) between each server site and the access client, the path information (logical distance) between each server site and the site to which the access client belongs may be utilized. For the network state information between each server site and the access client, the network state information between each server site and the access client may be utilized.

[0075] For example, when the IP address of the access client 4 is 192.168.0.1, the network state with this client is used as the result of the entire class C address block. Specifically, for the access from the network block of 192.168.0.0/24, the already inspected and obtained information of 192.168.0.1 is used. In this case, the memory space for the data to be stored can be reduced. In the present embodiment, the access client area is classified with the class B access (the network address block corresponding to A.B.0.0/16).

[0076] The network and information distribution server state information autonomously inspected and obtained by the NS server 11 will next be described.

(1) The network state information from each information distribution server site to the client are as follows:

- ASLn: distance between the logical networks by AS path (BGP path information is used);
- RTn: response time;
- RNN: number of router steps (number of router hops); and
- PLn: packet loss ratio.

(2) The state (network, server) information in each information distribution server site are as follows:
as the network state information in the site,

- CSn: congestion degree in the site;

- PSn: number of packets in the site; and
- ESn: packet error,

as the information distribution server state information,

- CPU_{nm}: CPU load ratio;
- IDLE_{nm}: CPU idle value;
- LINK_{nm}: number of connection links; and
- IO_{nm}: disk load ratio,

additionally, n: server site number, m: server number.

[0077] A method of using these network and server state information to determine the optimum server. The NS server 11 stores these various state information, and uses these information to determine the optimum server site and server. In the present embodiment, these are calculated using the following equation.

[0078] A selection method I corresponds to the selection method in the step S6 in Fig. 3B, and the site in which ASLn is minimum is fixed as the optimum site.

[0079] A selection method II corresponds to the selection method in the step S11 in Fig. 3D.

[0080] A network state value K1n between the server site and the client, a network state value K2n in the server site, and a server state value are obtained, and an optimum site judgment value Kn is obtained for each server site as shown by the following equation. The site number (n) in which the value is minimum is fixed as the optimum site. Subsequently, an optimum server judgment value S_{nm} is obtained using K3_{nm}. The optimum server judgment value is obtained from the optimum site judgment value and server state value. Therefore, by using the network state of the client site and server site, network state in the site, server state, and logical distance between the networks with weight coefficients A to O, the optimum server is generally judged. Additionally, here, the server in which S_{nm} is minimum is fixed as the optimum server.

Network state value:

$$K1n = RTn \cdot A + RNn \cdot B + PLn \cdot C$$

In-site network state value:

$$K2n = Csn \cdot D + PSn \cdot E + ESn \cdot F$$

Server state value:

$$K3nm = CPU_{nm} \cdot G + IDLE_{nm} \cdot H + LINK_{nm} \cdot I + IO_{nm} \cdot J$$

Optimum site judgment value:

$$Kn = K1n \cdot K + K2n \cdot L + ASLn \cdot M$$

Optimum server judgment value:

$$S_{nm} = Kn \cdot N + K3_{nm} \cdot O$$

[0081] Additionally, n: number of server sites, m: number of servers, A to O: weight coefficients.

[0082] Moreover, in the present embodiment, the response time, number of router steps (number of router hops), and packet loss ratio are used as the network state value, the congestion degree in the site, number of packets in the site, and packet error are used as the in-site network state value, and CPU load ratio, CPU idle value, number of connection links, and disk load ratio are used as the server state value, but with respect to the network state value, in-site network state value and server state value, at least one of the utilized information in the present embodiment may be utilized.

[0083] Moreover, the network state information in the site and the state information of the information distribution server are utilized as the state information in the server site, but either one may be used as the representative during the determination of the optimum server.

[0084] The function of the route server used in the present system will next be described with reference to Fig. 5. The current Internet/Intranet is said to be an environment in which routers are connected in a mesh manner. In the system, in order to distribute the desired packet to a specific place, a routing protocol is used, and the routing protocol called Border Gateway Protocol (BGP) is used in a wide network. In the BGP, each router is provided with a path table, and the packet is distributed to the specific place based on this path table.

[0085] In Fig. 5, the route server 10, each site router and BGP are logically connected as being peer-to-peer. Here, as one technique of measuring the distance between each site and the client, AS-path number is utilized. This AS-path number means the number of independent networks treated in the BGP, and the path information up to the desired AS number are mapped as the connection of AS numbers in the path information. Specifically, the path information between the server site and the client site can be known as the connection of AS numbers applied to the independent networks. The AS-path number indicates the frequency at which the independent networks are passed, and can be known from the connection of the AS numbers.

[0086] The route server can collect the path information viewed from each server site by obtaining the path information from the router 6 in each server site, and the server site closest to the access client (in the AS path) can be selected.

[0087] In the present embodiment, this route server is used to obtain the BGP distance between the access client and each server site, that is, the AS-path number, and the distance is used as one of the selection conditions for determining the optimum server site.

[0088] Moreover, in the present embodiment, used are the selection method I of obtaining the optimum site from the logical distance from the server site to the client and then determining the information distribution server in the optimum site in the predetermined method, and the selection method II of determining the optimum information distribution server from the logical distance, network state information and information distribution server state information. For the first access from the client, the selection method I is used to determine the optimum server, and for the second and subsequent accesses, the selection method II is used to determine the optimum server, but the present invention is not limited to this order, and the selection method I may constantly be used, or the selection method II may be used.

[0089] Moreover, when the client accesses the NS server, and the path information (logical distance) between the client and the server site is stored, the stored path information may be utilized to determine the optimum server.

[0090] Furthermore, the NS server, information distribution server and route server of the present invention can be realized by the information processor with a constitution equal to that of the computer shown in Fig. 9. A computer 901 of Fig. 9 is constituted of a CPU 902 for reading a program to perform an actual processing, RAM 903 used by the CPU 902 during the processing, hard disk (hereinafter referred to as HD) 904 as the storage medium for supplying a program code, floppy disk (FD) 905, communication interface 906 used during connection to the network, and the like.

The network status server control program and information distribution system control program of the present invention may be performed by the computer with the constitution equal to that of Fig. 9 by the program installed from the outside. In this case, as shown in Fig. 10, needless to say, the object of the present invention is achieved by providing the system or the device with a recording medium 1002 in which the software program code for realizing the function of the above-described embodiment is recorded, and reading and executing the program code stored in the storage medium by the computer (or CPU or MPU) of the system or the device.

[0091] In this case, the program code itself read from the storage medium realizes the inventive function of the present invention, and the storage medium for storing the program code constitutes the present invention.

[0092] As the storage medium for supplying the program code, for example, magnetic disks such as a floppy disk and hard disk, optical disks, optical magnetic disks, CD-ROM, CD-R, DVD-ROM, DVD-RAM, magnetic tapes, memory cards, ROM, and the like can be used.

[0093] Moreover, it goes without saying that the present invention can be applied to a case in which the program is distributed from the storage medium for recording the program code of the software for realizing the function of the above-described embodiment via the communication line of the personal computer communication, and the like.

[0094] Fig. 11 is a diagram showing one example of the memory map of storage media such as the CD-ROM for supplying the program code. An area 1101 for storing directory information indicates the positions of an area 1102 for storing an install program, and an area 1103 for storing control programs such as a network status server control program and an information distribution system control program. When the network status server control program or the information distribution system control program of the present invention is installed in the computer, the install program stored in the area 1102 is loaded to the system, and executed by the CPU. Subsequently, the install program executed by the CPU reads the network status server control program or the information distribution system control program from the area 1103 for storing the control programs such as the network status server control program and information distribution system control program and stores the program in the hard disk.

[0095] When the present invention is applied to the storage medium, the storage medium stores the program code

corresponding to the flowchart described above.

[0096] Moreover, it goes without saying that, as shown in Fig. 10, the object of the present invention is achieved by transmitting the program code of the software for realizing various functions in the above-described embodiment from a transmission device 1004, for example HTTP server or FTP server, receiving the code via a network 1005, public line or radio, and executing the program code by the computer (or CPU, MPU) of the system or device.

[0097] In this case, since the program code itself transmitted from the transmission device realizes the function in the above-described embodiment, the transmission device for transmitting the program code constitutes the present invention.

[0098] Moreover, the present invention can be applied not only to the realization of the function of the above-described embodiment by executing the program code read by the computer but also to the realization of the function of the above-described embodiment by performing a part or the whole of the actual processing by the OS, and the like operating on the computer based on the instruction of the program code.

[0099] As described above, according to the present embodiment, there is an effect that the access load between the dispersed/arranged web server sites can be leveled.

[0100] Moreover, the access between the dispersed/arranged web server sites is leveled, and the access load can effectively be leveled even between the web servers in each web server site.

[0101] Furthermore, in order to determine the optimum web server at a high speed, when the client has a first access, the optimum web server site is determined using the path information (logical distance) at a high speed in the first determination method. In the subsequent access from the client, the optimum web server can be determined at a high speed by using not only the path information but also the network state information and the state information in the web server site in the second determination method to determine the optimum web server, which produces an effect that the web server can respond to the access client at a high speed.

[0102] Additionally, since the network state information and server state information necessary for recognizing the optimum web server are automatically obtained and stored in the NS server, or the NS server obtains and stores the information, and the optimum web server is generally determined by the information, there is an effect that the accessed web server can determine the optimum web server at a high speed.

[0103] Moreover, the client is switched to the optimum web server without adding the separate device or requiring any active operation, and all the web servers in the dispersed/arranged web server sites can accept the access, the concentration of the accesses can effectively be eliminated.

[0104] Further, the computer program for carrying out the above disclosed method can be obtained in electronic form for example by downloading the code over a network such as the internet. Thus in accordance with another aspect of the present invention there is provided an electrical signal carrying processor implementable instructions for controlling a processor to carry out the method as hereinbefore described.

Claims

1. A network status server comprising:

logical distance obtaining means for obtaining respective logical distances between respective sites provided with dispersed/arranged information distribution servers and an accessed client; and
site determination means for determining an optimum site from said respective sites based on said logical distance.

2. A network status server comprising:

logical distance obtaining means for obtaining respective logical distances between respective sites provided with dispersed/arranged information distribution servers and an accessed client;
collection means for collecting network state information with said client, and state information in said respective sites; and
server determination means for determining the optimum information distribution server from said dispersed/arranged information distribution servers based on the logical distance obtained by said logical distance obtaining means, the network state information between said respective sites and said client, and the state information in said site collected by said collection means.

3. A network status server comprising:

logical distance obtaining means for obtaining respective logical distances between respective sites provided

with dispersed/arranged information distribution servers and an accessed client;
collection means for collecting network state information between said respective sites and said client, and
state information in said site;
site determination means for, when it is judged that the access from said client is a first access, determining
the optimum site from the respective sites based on the logical distance obtained by said logical distance
obtaining means;
first server determination means for determining the optimum information distribution server from the informa-
tion distribution servers in the optimum site determined by said site determination means based on a prede-
termined condition; and
second server determination means for, when it is judged that the access from the client is not the first access,
determining the optimum information distribution server from said dispersed/arranged information distribution
servers based on the logical distance obtained from said logical distance obtaining means, the network state
information between said respective sites and said client, and the state information in said site collected by
said collection means.

4. An information distribution system comprising: dispersed/arranged information distribution servers; and a network
status server,
said information distribution server comprising:

inquiry means for inquiring of said network status server about the optimum information distribution server
among said dispersed/arranged information distribution servers in response to an accessing client;
response reception means for receiving a response from said network status server; and
indication means for indicating said optimum information distribution server to said accessing client,

said network status server comprising:

inquiry reception means for receiving the inquiry from said inquiry means;
logical distance obtaining means for obtaining respective distances between said respective sites provided
with the dispersed/arranged information distribution servers and said client;
site determination means for determining the optimum site from the respective sites based on said obtained
logical distance;
server determination means for determining the optimum information distribution server from the information
distribution servers in the optimum site determined by said site determination means based on a predetermined
condition; and
response transmission means for returning the server determined by said server determination means as the
optimum server to said dispersed/arranged information distribution servers having transmitted the inquiry.

5. An information distribution system comprising: dispersed/arranged information distribution servers; and a network
status server,

said information distribution server comprising:

inquiry means for inquiring of said network status server about the optimum information distribution server
among said dispersed/arranged information distribution servers in response to an accessing client;
response reception means for receiving a response from said network status server; and
indication means for indicating said optimum information distribution server to said accessing client,

said network status server comprising:

inquiry reception means for receiving the inquiry from said inquiry means;
logical distance obtaining means for obtaining respective distances between respective sites provided with
said dispersed/arranged information distribution servers and said client;
collection means for collecting network state information between said respective sites and said client, and
state information in said site;
server determination means for determining the optimum information distribution server from said dispersed/
arranged information distribution servers based on the logical distance obtained by said logical distance ob-
taining means, the network state information between said respective sites and said client, and said state
information in the site collected by said collection means; and
response transmission means for returning the server determined by said server determination means as said

optimum server to said dispersed/arranged information distribution servers having transmitted the inquiry.

6. An information distribution system comprising: dispersed/arranged information distribution servers; and a network status server,

said information distribution server comprising:

inquiry means for inquiring of said network status server about the optimum information distribution server among said dispersed/arranged information distribution servers in response to an accessing client;
response reception means for receiving a response from said network status server; and
indication means for indicating the optimum information distribution server to said client,

said network status server comprising:

inquiry reception means for receiving the inquiry from said inquiry means;
logical distance obtaining means for obtaining respective distances between respective sites provided with said dispersed/arranged information distribution servers and the accessing client;
collection means for collecting network state information between said respective sites and said client, and state information in the respective sites;
site determination means for, when it is judged that the access from said client is a first access, determining the optimum site from said respective sites based on the logical distance obtained by said logical distance obtaining means;
first server determination means for determining the optimum information distribution server from the information distribution servers in said respective sites based on a predetermined condition;
second server determination means for, when it is judged that the access from said client is not the first access, determining the optimum information distribution server from said dispersed/arranged information distribution servers based on the logical distance obtained from said logical distance obtaining means, the network state information between said respective sites and said client, and the state information in said respective sites collected by said collection means; and
response transmission means for returning the server determined by said first server determination means or said second server determination means as said optimum server to said dispersed/arranged information distribution servers having transmitted the inquiry.

7. A network status server control method of controlling a network status server, the method comprising:

a logical distance obtaining step of obtaining respective logical distances between respective sites provided with dispersed/arranged information distribution servers and an accessed client; and
site determining step of determining an optimum site from said respective sites based on said obtained logical distance.

8. The network status server control method according to claim 7 wherein said logical distance obtaining step comprises obtaining said logical distance from a route server comprising means for obtaining the logical distance to a predetermined network address from path information between the sites.

9. The network status server control method according to claim 8 wherein said network status server comprises path information storage means,

said network status server control method further comprises a path information storing step of storing the path information between said respective sites and said client, and
when said path information is stored in said path information storage means, said logical distance obtaining step comprises obtaining said logical distance from the path information stored in said path information storage means.

10. The network status server control method according to claim 9 wherein when there is no access from said client for a predetermined period, said path information storing step comprises discarding the path information between said respective sites and the client stored in said path information storage means.

11. The network status server control method according to claim 7 wherein said site determining step comprises determining the site in which the respective logical distances between said respective sites and said client are min-

imum as said optimum site.

12. The network status server control method according to claim 7, further comprising a server determining step of determining the optimum information distribution server from the information distribution servers in the optimum site determined in said site determining step based on a predetermined condition.

13. The network status server control method according to claim 12, further comprising a state information collecting step of collecting state information in the site provided with said dispersed/arranged information distribution servers, wherein said server determining step comprises determining the optimum information distribution server using the collected state information in said optimum site as a condition.

14. The network status server control method according to claim 13 wherein said network status server comprises state information storage means,

said network status server control method further comprises a state information storing step of storing the collected state information in said site into the state information storage means, and when the state information in said optimum site is stored in said state information storage means, said server determining step comprises determining the optimum information distribution server using the stored state information in said optimum site as the condition.

15. The network status server control method according to claim 14, further comprising a state information discarding step of discarding the state information in the site stored in said state information storage means after a predetermined period elapses.

16. The network status server control method according to claim 14 wherein said state information storing step comprises approximating and storing said state information in the site by a predetermined approximation equation.

17. The network status server control method according to claim 13 wherein said state information in the site comprises at least one information of said network state information in the site and said state information of the information distribution server in the site.

18. A network status server control method for controlling a network status server, the method comprising:

a logical distance obtaining step of obtaining respective logical distances between respective sites provided with dispersed/arranged information distribution servers and an accessed client;
a collecting step of collecting network state information between said respective sites and said client, and state information in said site; and
a server determining step of determining the optimum information distribution server from said information distribution servers based on said obtained logical distance, the network state information between said respective sites and said client, and the state information in said site collected in said collecting step.

19. The network status server control method according to claim 18 wherein said logical distance obtaining step comprises obtaining said logical distance from a route server comprising means for obtaining the logical distance to a predetermined network address from path information between the sites.

20. The network status server control method according to claim 19 wherein said network status server comprises path information storage means,

said network status server control method further comprises a path information storing step of storing the path information between said respective sites and said client into said path information storage means, and when said path information is stored in said path information storage means, said logical distance obtaining step comprises obtaining said logical distance from the path information stored in said path information storing step.

21. The network status server control method according to claim 20 wherein when there is no access from said client for a predetermined period, said path information storing step comprises discarding the path information between said respective sites and the client stored in said path information storage means.

22. The network status server control method according to claim 18 wherein said network status server comprises state information storage means,

said network status server control method further comprises a state information storing step of storing the network state information between said respective sites and said client, and the state information in said site into the state information storage means, and

when said state information is stored in said state information storage means, said collecting step comprises collecting said state information from said state information stored in said state information storage means.

23. The network status server control method according to claim 22 wherein when there is no access from said client for a predetermined period, said state information storing step comprises discarding said network state information between the respective sites and the client stored in said state information storage means.

24. The network status server control method according to claim 22, further comprising a state information discarding step of discarding the state information in the site stored in said state information storage means after the predetermined period elapses.

25. The network status server control method according to claim 22 wherein said state information storing step comprises approximating and storing said state information in the site by a predetermined approximation equation.

26. The network status server control method according to claim 18 wherein said state information in the respective sites comprises at least one of said network state information in the respective sites and said state information of the dispersed/arranged information distribution servers.

27. The network status server control method according to claim 18 wherein said network status server comprises path information storage means and state information storage means,

said network status server control method further comprises: a path information storing step of storing path information between said respective sites and said client into the path information storage means; and a state information storing step of storing network state information between said respective sites and said client, and state information in said respective sites into the state information storage means, said server determining step comprises determining the information distribution server in which an optimum server judgment value S_{nm} in the following equations satisfies a predetermined condition as the optimum information distribution server based on said stored path information, said state information and a predetermined weight coefficient;

- network state value:

$$K1n = RTn \cdot A + RNn \cdot B + PLn \cdot C$$

- in-site network state value:

$$K2n = Csn \cdot D + PSn \cdot E + ESn \cdot F$$

- server state value:

$$K3nm = CPU_{nm} \cdot G + IDLE_{nm} \cdot H + LINK_{nm} \cdot I + IOn_{m} \cdot J$$

- optimum site judgment value:

$$K_n = K1n \cdot K + K2n \cdot L + AsLn \cdot M$$

- optimum server judgment value: $S_{nm} = K_n \cdot N + K3nm \cdot O$ (additionally, n: number of sites, m: number of servers, A to O: weight coefficients, and for respective symbols, $ASLn$: logical distance between the net-

works by AS path (using the path information of BGP), RTn: response time, RNn: number of router steps (number of router hops), PLn: packet loss ratio, CSn: congestion degree in the site, PSn: number of packets in the site, ESn: number of packet errors, CPUnm: CPU load ratio, IDLEnm: CPU idle value, LINKnm: number of connection links, IOnm: disk load ratio).

28. A network status server control method comprising:

a logical distance obtaining step of obtaining respective logical distances between respective sites provided with dispersed/arranged information distribution servers and an accessed client;
a collecting step of collecting network state information between said respective sites and said client, and state information in said respective sites;
a site determining step of, when it is judged that the access from said client is a first access, determining the optimum site from said respective sites based on said obtained logical distance;
a first server determining step of determining the optimum information distribution server from the information distribution servers in said optimum site based on a predetermined condition; and
a second server determining step of, when it is judged that the access from said client is not the first access, determining the optimum information distribution server from said dispersed/arranged information distribution servers based on said obtained logical distance, said network state information between said respective sites provided with the dispersed/arranged information distribution servers and said client, and said state information in the site provided with the dispersed/arranged information distribution servers collected in said collecting step.

29. The network status server control method according to claim 28 wherein said logical distance obtaining step comprises obtaining said logical distance from a route server comprising means for obtaining the logical distance to a predetermined network address from path information between the sites.

30. The network status server control method according to claim 29 wherein said network status server comprises path information storage means,

said network status server control method further comprises a path information storing step of storing path information between said respective sites provided with the dispersed/arranged information distribution servers and said client into said path information storage means, and
when said path information is stored in said path information storage means, said logical distance obtaining step comprises obtaining said logical distance from said path information stored in the path information storage means.

31. The network status server control method according to claim 30 wherein when there is no access from said client for a predetermined period, said path information storing step comprises discarding said path information between the respective sites and the client stored in said path information storage means.

32. The network status server control method according to claim 28 wherein said network status server comprises state information storage means,

said network status server control method comprises a state information storing step of storing network state information between said respective sites provided with the dispersed/arranged information distribution servers and said client, and state information in said site provided with the dispersed/arranged information distribution servers into said state information storage means, and
when said state information is stored in said state information storage means, said collecting step comprises collecting said state information from said state information stored in said state information storage means.

33. The network status server control method according to claim 32 wherein when there is no access from said client for a predetermined period, said state information storing step comprises discarding said network state information between the respective sites and the client stored in said state information storage means.

34. The network status server control method according to claim 32, further comprising a state information discarding step of discarding the state information in the site stored in said state information storage means after the predetermined period elapses.

35. The network status server control method according to claim 32 wherein said state information storing step comprises approximating and storing said state information in the site by a predetermined approximation equation.

36. The network status server control method according to claim 28 wherein said network status server comprises path information storage means and state information storage means,

said network status server control method further comprises: a path information storing step of storing path information between said respective sites provided with the dispersed/arranged information distribution servers and said client into said path information storage means; and

a state information storing step of storing network state information between said respective sites provided with the dispersed/arranged information distribution servers and said client, and state information in said site provided with the dispersed/arranged information distribution servers into said state information storage means, and

said second server determining step comprises determining the information distribution server in which an optimum server judgment value S_{nm} in the following equations satisfies a predetermined condition as the optimum information distribution server based on said stored path information, said state information and a predetermined weight coefficient;

- network state value:

$$K1n = RTn \cdot A + RNn \cdot B + PLn \cdot C$$

- in-site network state value:

$$K2n = CSn \cdot D + PSn \cdot E + ESn \cdot F$$

- server state value:

$$K3nm = CPU_{nm} \cdot G + IDLE_{nm} \cdot H + LINK_{nm} \cdot I + IO_{nm} \cdot J$$

- optimum site judgment value:

$$K_n = K1n \cdot K + K2n \cdot L + ASLn \cdot M$$

- optimum server judgment value:

$$S_{nm} = K_n \cdot N + K3nm \cdot O$$

(additionally, n: number of sites, m: number of servers, A to O: weight coefficients, and for respective symbols, ASLn: logical distance between the networks by AS path (using the path information of BGP), RTn: response time, RNn: number of router steps (number of router hops), PLn: packet loss ratio, CSn: congestion degree in the site, PSn: number of packets in the site, ESn: number of packet errors, CPU_{nm}: CPU load ratio, IDLE_{nm}: CPU idle value, LINK_{nm}: number of connection links, and IO_{nm}: disk load ratio).

37. An information distribution system control method for controlling an information distribution system comprising: dispersed/arranged information distribution servers; and a network status server,

said method comprising: in said information distribution server, an inquiring step of inquiring of said network status server about the optimum information distribution server among said information distribution servers in response to an accessing client; a response receiving step of receiving a response from said network status server; and an indicating step of indicating said optimum information distribution server to said accessing client, in said network status server,

an inquiry receiving step of receiving the inquiry from said inquiring step;
a logical distance obtaining step of obtaining respective distances between said respective sites provided with the dispersed/arranged information distribution servers and said client;
a site determining step of determining the optimum site from said respective sites provided with the dispersed/arranged information distribution servers based on said obtained logical distance;
a server determining step of determining the optimum information distribution server from the information distribution servers in the optimum site determined by said site determining step based on a predetermined condition; and
a response transmitting step of returning the server determined in said server determining step as said optimum server to said information distribution server having transmitted the inquiry.

38. The information distribution system control method according to claim 37 wherein said information distribution system further comprises a route server,

said control method comprises, in said route server a logical distance calculating step of obtaining a logical distance to a predetermined network address from the path information between the sites, and said logical distance obtaining step comprises obtaining said logical distance calculated in said logical distance calculating step.

39. The information distribution system control method according to claim 37, further comprising: in said information distribution server,

a step of collecting network state information between the site to which the information distribution server belongs and said accessing client; and
a network state transmitting step of transmitting said collected network state information to the network status server.

40. The information distribution system control method according to claim 37, further comprising: in said information distribution server, a step of collecting network state information in the site to which the information distribution server belongs, and the state information of the information distribution server; and

an in-site information transmitting step of transmitting said collected network state information in the site, and the state information of the information distribution server to the network status server.

41. An information distribution system control method for controlling an information distribution system comprising: dispersed/arranged information distribution servers; and a network status server,

said method comprising: in said information distribution server,
an inquiring step of inquiring of said network status server about the optimum information distribution server among said information distribution servers in response to an accessing client;
a response receiving step of receiving a response from said network status server; and
an indicating step of indicating said optimum information distribution server to said client, and in said network status server,

an inquiry receiving step of receiving the inquiry from said inquiring step;
a logical distance obtaining step of obtaining respective logical distances between said respective sites and said client;
a collecting step of collecting network state information between said respective sites and said client, and state information in said respective sites;
a server determining step of determining the optimum information distribution server from said dispersed/arranged information distribution servers based on the logical distance obtained in said logical distance obtaining step, the network state information between said respective sites and said client, and said state information in the site collected in said collecting step; and
a response transmitting step of returning the server determined in said server determining step as said optimum server to said dispersed/arranged information distribution servers having transmitted the inquiry.

42. The information distribution system control method according to claim 41 wherein said information distribution system comprises a route server,

said control method comprises, in said route server, a logical distance calculating step of obtaining a logical

distance to a predetermined network address from the path information between the sites, and said logical distance obtaining step comprises obtaining said logical distance calculated in said logical distance calculating step.

5 43. The information distribution system control method according to claim 41, further comprising: in said information distribution server,

a step of collecting network state information between the site to which the information distribution server belongs and said accessing client; and

10 a network state transmitting step of transmitting said collected network state information to the network status server.

44. The information distribution system control method according to claim 41, further comprising: in said information distribution server,

15 a step of collecting network state information of the site to which the information distribution server belongs, and the state information of the information distribution server; and

an in-site information transmitting step of transmitting said collected network state information in the site, and the state information of the information distribution server to the network status server.

20 45. An information distribution system control method for controlling an information distribution system comprising: dispersed/arranged information distribution servers; and a network status server,

said control method comprising: in said information distribution server,

25 an inquiring step of inquiring of said network status server about the optimum information distribution server among said information distribution servers in response to an accessing client;

a response receiving step of receiving a response from said network status server; and

an indicating step of indicating said optimum information distribution server to said client, in said network status server,

30 an inquiry receiving step of receiving the inquiry from said inquiring step;

a logical distance obtaining step of obtaining respective distances between said respective sites and said client;

a collecting step of collecting network state information between said respective sites and said client, and state information in said respective sites;

35 a site determining step of, when it is judged that the access from said client is a first access, determining the optimum site from said respective sites provided with the dispersed/arranged information distribution servers based on the logical distance obtained in said logical distance obtaining step;

a first server determining step of determining the optimum information distribution server from the information distribution servers in the optimum site determined in said site determining step based on a predetermined condition;

40 a second server determining step of, when it is judged that the access from said client is not the first access, determining the optimum information distribution server from said information distribution servers based on the logical distance obtained in said logical distance obtaining step, the network state information between said respective sites provided with the dispersed/arranged information distribution servers and said client, and said state information in the site provided with the dispersed/arranged information distribution servers collected

45 in said collecting step; and

a response transmitting step of returning the server determined in said first server determining step or said second server determining step as said optimum server to said dispersed/arranged information distribution servers having transmitted the inquiry.

50 46. The information distribution system control method according to claim 45 wherein said information distribution system further comprises a route server,

said control method comprises, in said route server, a logical distance calculating step of obtaining a logical distance to a predetermined network address from the path information between the sites, and

55 said logical distance obtaining step comprises obtaining said logical distance calculated in said logical distance calculating step.

47. The information distribution system control method according to claim 45, further comprising: in said information

distribution server,

a step of collecting network state information between the site to which the information distribution server belongs and said accessing client; and

a network state transmitting step of transmitting said collected network state information to the network status server.

48. The information distribution system control method according to claim 45, further comprising: in said information distribution server,

a step of collecting network state information of the site to which the information distribution server belongs, and the state information of the information distribution server; and

an in-site information transmitting step of transmitting said collected network state information in the site, and the state information of the information distribution server to the network status server.

49. A computer readable medium storing a computer program element executing any one of methods as claimed in claims 7 to 48.

50. A computer program comprising processor implementable instructions for carrying out a method as claimed in any one of claims 7 to 48.

51. An electrical signal carrying processor implementable instructions for controlling a processor to carry out the method of any one of claims 7 to 48.

FIG. 1

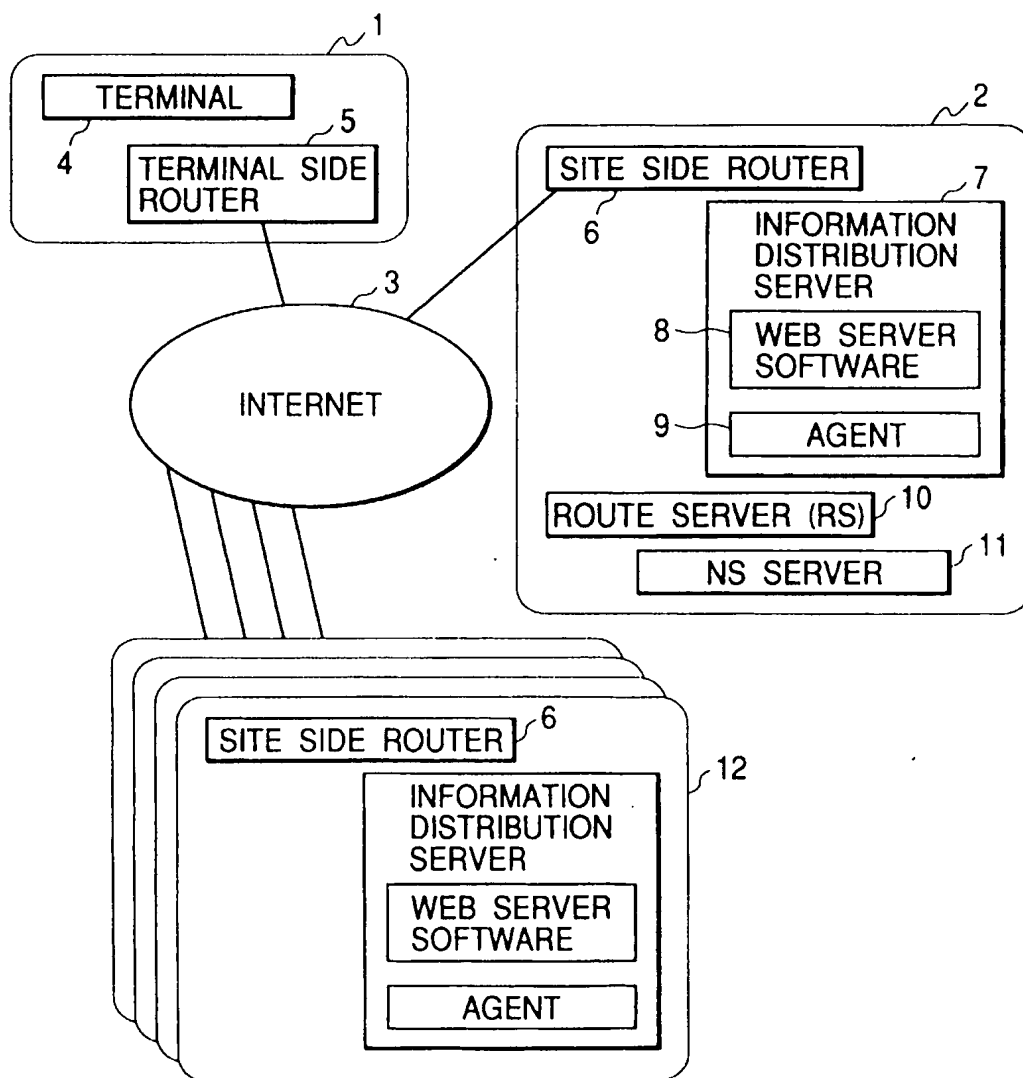


FIG. 2

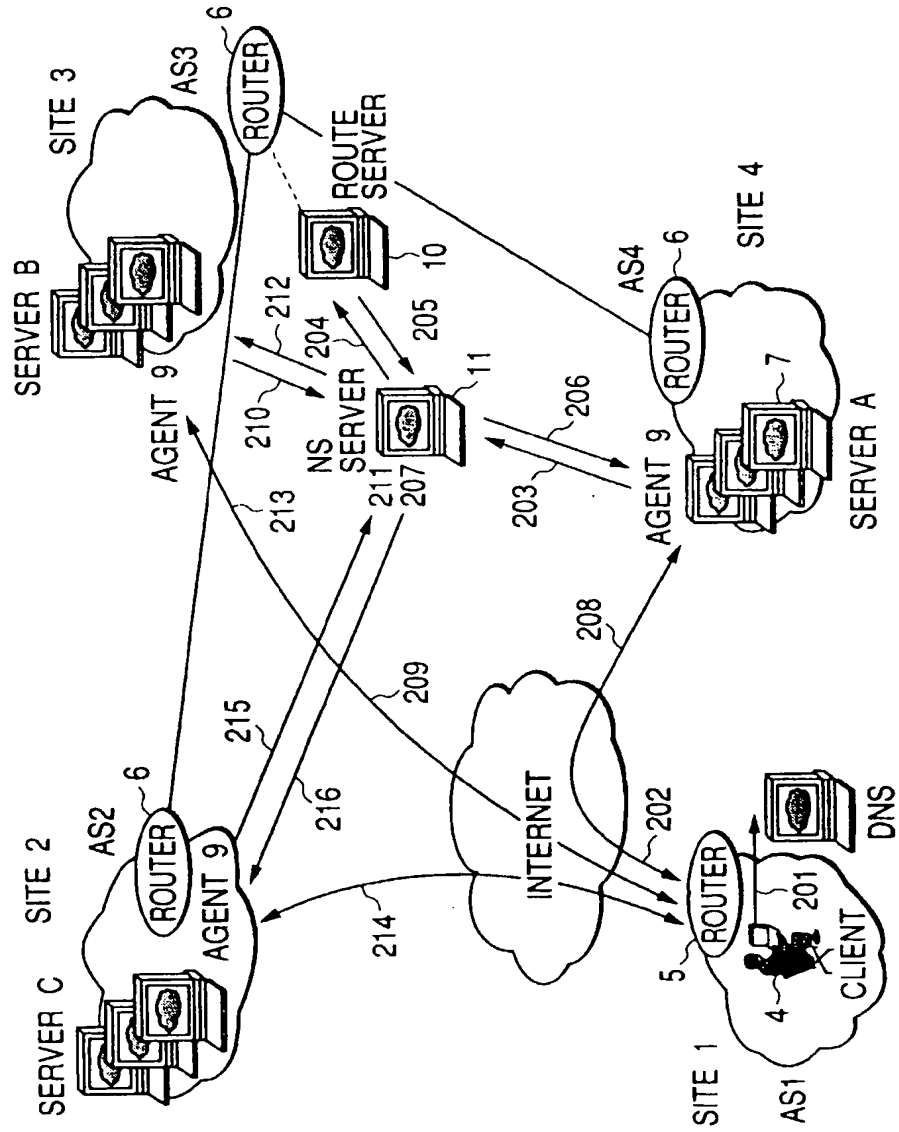


FIG. 3A

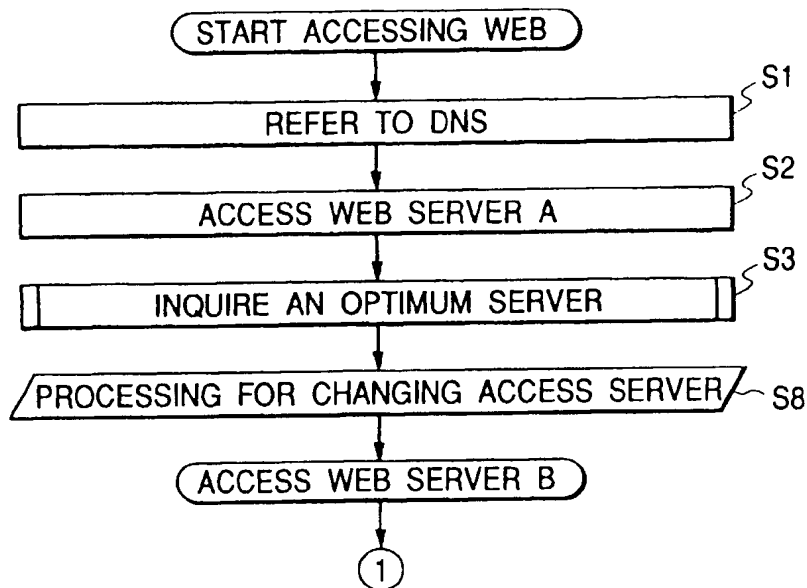


FIG. 3B

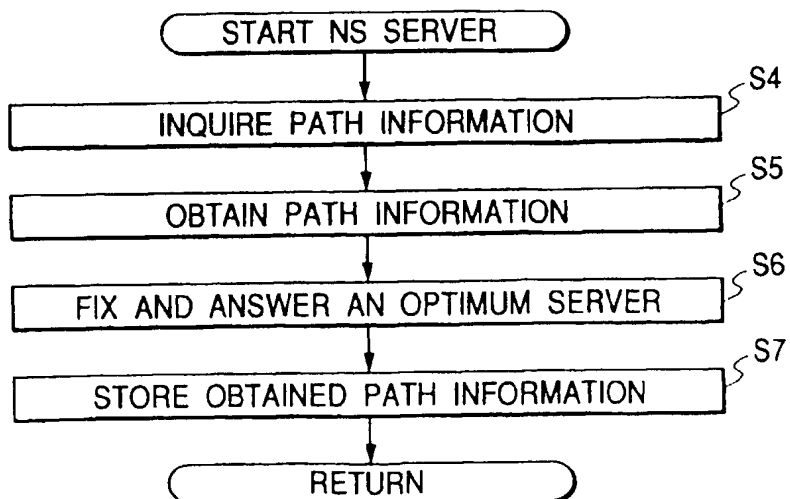


FIG. 3C

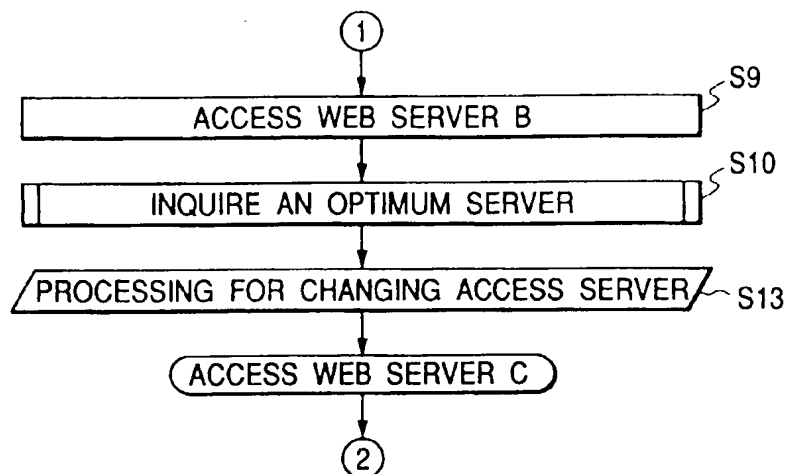


FIG. 3D

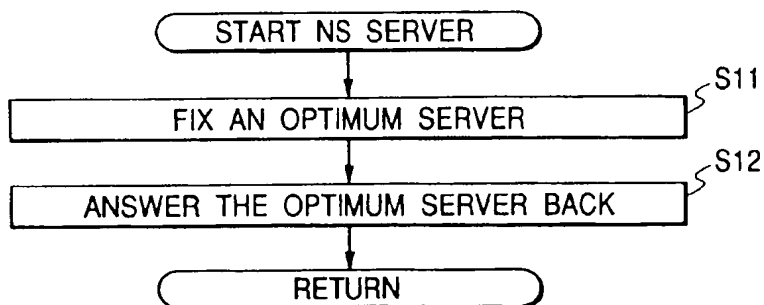


FIG. 3E

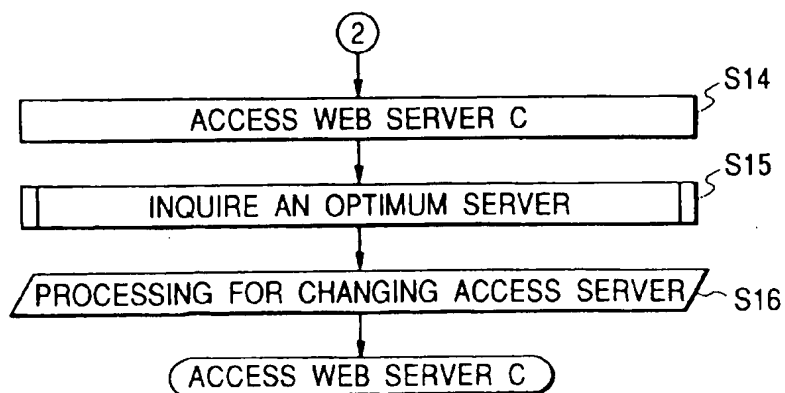


FIG. 4

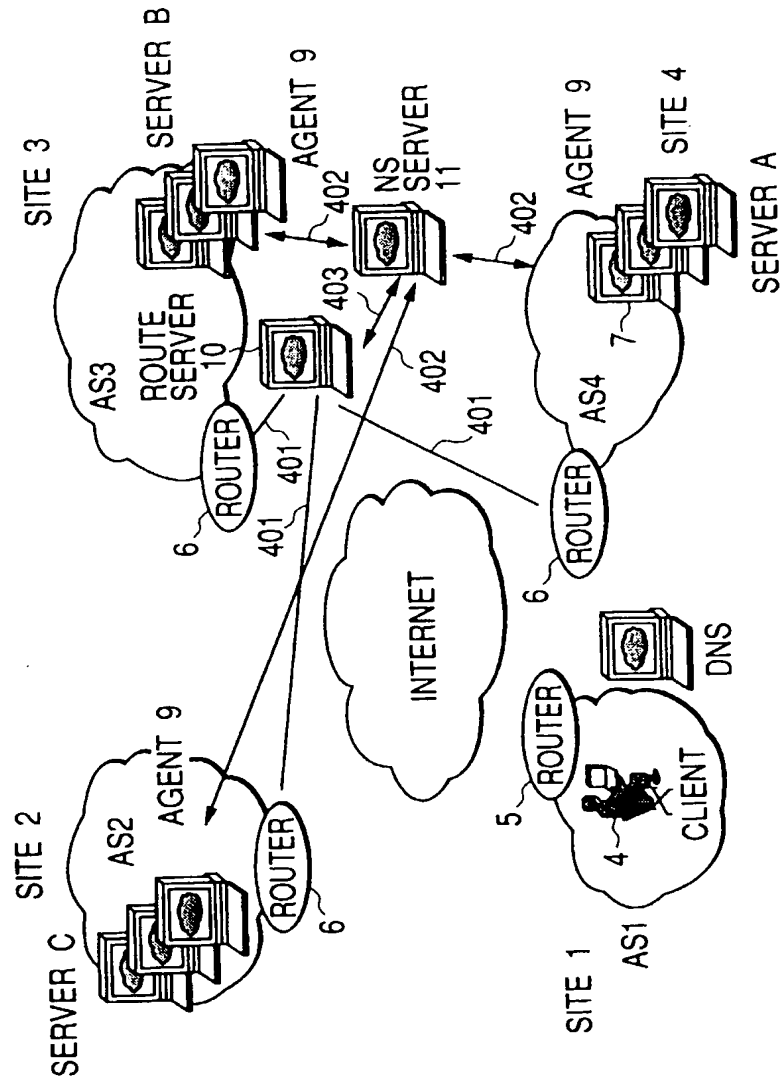


FIG. 5

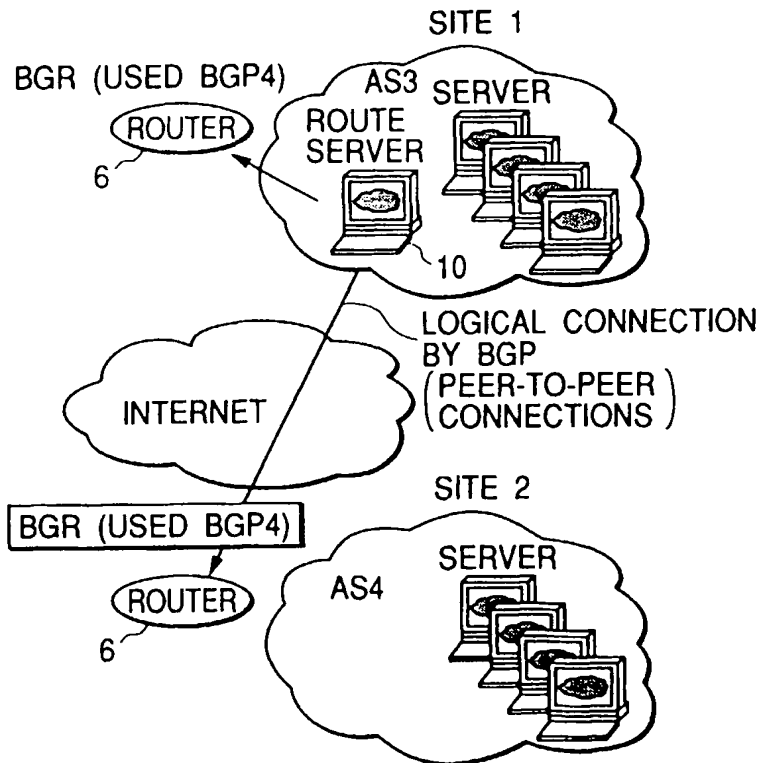


FIG. 6

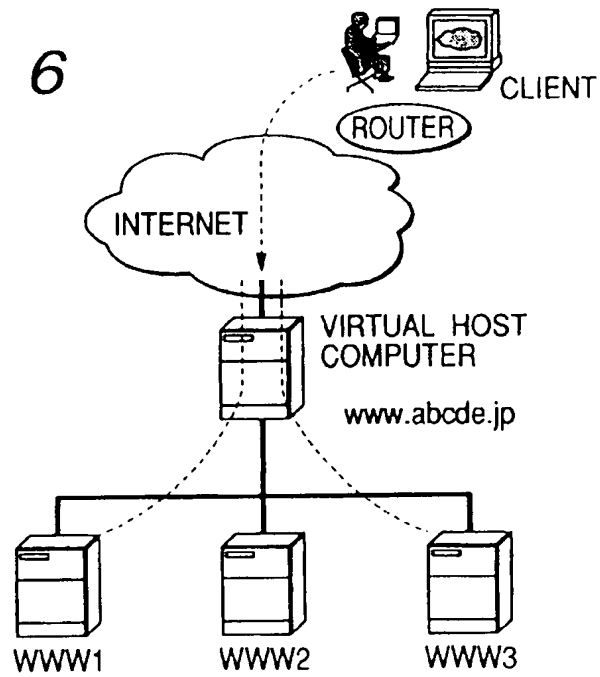


FIG. 7

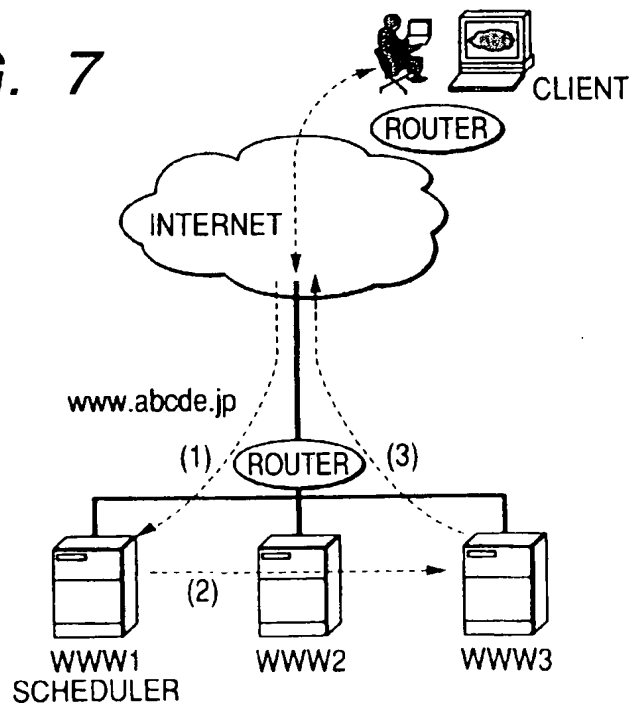


FIG. 8

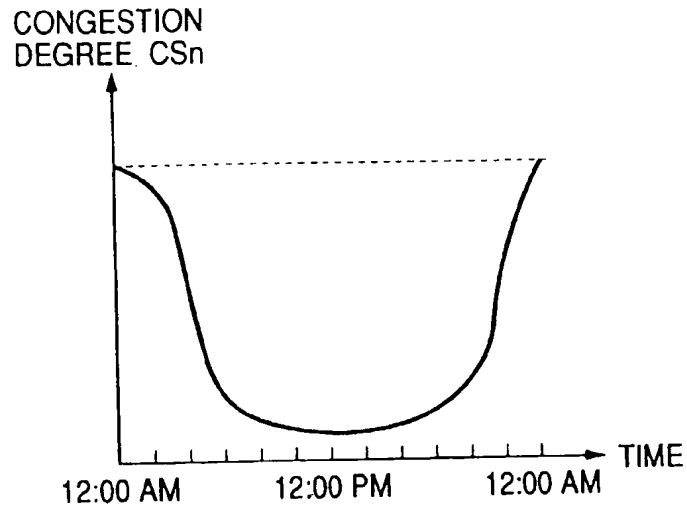


FIG. 9

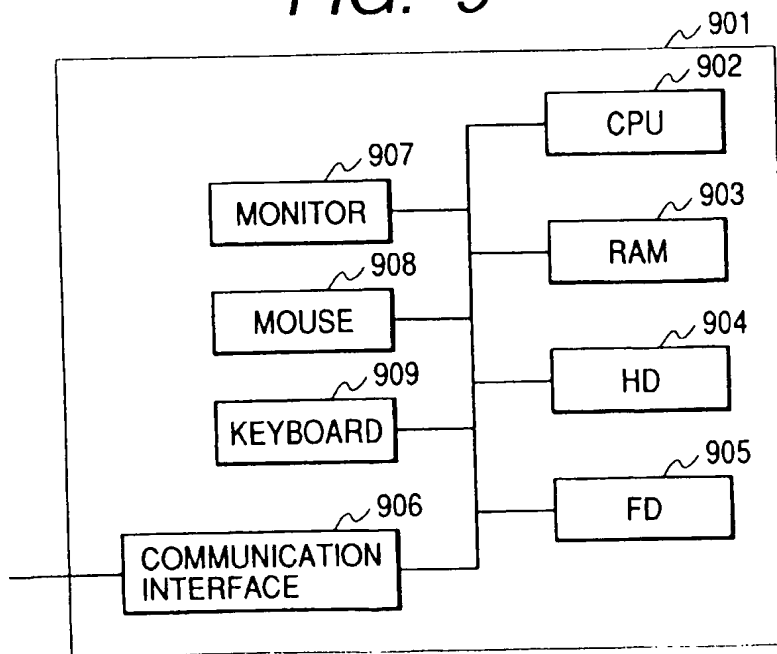


FIG. 10

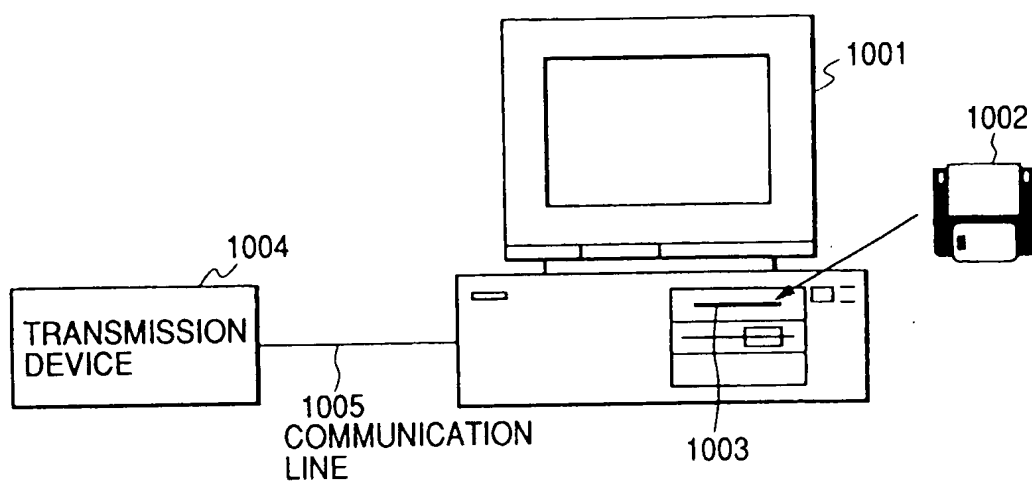


FIG. 11

